(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
9 January 2003 (09.01.2003)

PCT

(10) International Publication Number
WO 03/002604 A2

(51) International Patent Classification[7]: C07K 14/705

(21) International Application Number: PCT/EP02/07021

(22) International Filing Date: 25 June 2002 (25.06.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/300,944 26 June 2001 (26.06.2001) US

(71) Applicant *(for all designated States except AT, US)*: **NOVARTIS AG** [CH/CH]; Lichtstrasse 35, CH-4056 Basel (CH).

(71) Applicant *(for AT only)*: **NOVARTIS-ERFINDUNGEN VERWALTUNGSGESELLSCHAFT M.B.H.** [AT/AT]; Brunner Strasse 59, A-1230 Vienna (AT).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **BHATIA, Umesh** [US/US]; 5212 Union Avenue, San Jose, CA 95124 (US). **JONES, Carol, Elizabeth** [GB/GB]; Novartis Horsham Research Centre, Wimblehurst Road,, Horsham, West Sussex RH12 5AB (GB). **BOUHELAL, Rochdi** [FR/FR]; 2a rue des Fleurs, F-68200 Buschwiller (FR). **SEUWEN, Klaus** [DE/FR]; 11, rue des Lilas, F-68730 Michelbach le Bas (FR). **TENAILLON, Laurent** [FR/FR]; 13, rue des 7 clochers, F-68440 Dietwiller (FR).

(74) Agent: **GROS, Florent**; Novartis AG, Corporate Intellectual Property, Patent & Trademark Department, CH-4002 Basel (CH).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LT, LU, LV, MA, MD, MK, MN, MX, NO, NZ, OM, PH, PL, PT, RO, RU, SE, SG, SI, SK, TJ, TM, TN, TR, TT, UA, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 03/002604 A2

(54) Title: NOVEL G PROTEIN-COUPLED RECEPTORS AND DNA SEQUENCES THEREOF

(57) Abstract: This invention relates to newly identified polypeptides and polynucleotides encoding respiratory chemokine receptor polypeptides and polynucleotides, to their use in diagnosis and in identifying compounds that may be agonists and antagonists that are potentially useful in therapy, and to production of such polypeptides and polynucleotides, belonging to the class of G protein-coupled receptors.

Novel G Protein-Coupled Receptors and DNA Sequences Thereof

Field of the Invention

This invention relates to newly identified polypeptides and polynucleotides encoding such polypeptides, to their use in diagnosis and in identifying compounds that may be agonists, antagonists that are potentially useful in therapy, and to production of such polypeptides and polynucleotides, belonging to the class of G protein-coupled receptors.

Background of the Invention

The drug discovery process is currently undergoing a fundamental revolution as it embraces "functional genomics", that is, high throughput genome- or gene-based biology. This approach as a means to identify genes and gene products as therapeutic targets is rapidly superceding earlier approaches based on "positional cloning". A phenotype, that is a biological function or genetic disease, would be identified and this would then be tracked back to the responsible gene, based on its genetic map position.

Functional genomics relies heavily on high-throughput DNA sequencing technologies and the various tools of bioinformatics to identify gene sequences of potential interest from the many molecular biology databases now available. There is a continuing need to identify and characterise further genes and their related polypeptides/proteins, as targets for drug discovery.

Summary of the Invention

The present invention relates to in particular respiratory chemokine receptor polypeptides and polynucleotides, recombinant materials and methods for their production. Such polypeptides and polynucleotides are of interest in relation to methods of treatment of certain diseases, including, but not limited to,asthma, chronic obstructive pulmonary disease, emphysema, chronic bronchitis, acute respiratory distress syndrome, cough and acute bronchitis, hereinafter referred to as " diseases of the invention". In a further aspect, the invention relates to methods for identifying agonists and antagonists (e.g., inhibitors) using the materials provided by the invention, and treating conditions associated with monocyte/macrophage migration/activation, airway remodeling, airway fibrosis, regulation of epithelial differentiation, regulation of mucus hypersecretion, regulation of mucocilliary

clearance, regulation of inflammation, modulation of neutrophil, T-cell and eosinophil migration and/or activation and regulation of epithelial cell or mast cell activation with the identified compounds. In a still further aspect, the invention relates to diagnostic assays for detecting diseases associated with inappropriate activity/levels of monocyte/macrophage migration/activation, airway remodeling, airway fibrosis, regulation of epithelial differentiation, regulation of mucus hypersecretion, regulation of mucocilliary clearance, regulation of inflammation, modulation of neutrophil, T-cell and eosinophil migration and/or activation and regulation of epithelial cell or mast cell activation.

Description of the Invention

In a first aspect, the present invention relates to respiratory chemokine receptor polypeptides.

Such polypeptides include:

(a) an isolated polypeptide encoded by a polynucleotide comprising the sequence of SEQ ID NO 1 or SEQ ID NO 5;

(b) an isolated polypeptide comprising a polypeptide sequence having at least 95%, 96%, 97%, 98%, or 99% identity to the polypeptide sequence of SEQ ID NO 2 or SEQ ID NO 6;

(c) an isolated polypeptide comprising the polypeptide sequence of SEQ ID NO NO 2 or SEQ ID NO 6;

(d) an isolated polypeptide having at least 95%, 96%, 97%, 98%, or 99% identity to the polypeptide sequence of SEQ ID NO 2 or SEQ ID NO 6;

(e) the polypeptide sequence of SEQ ID NO 2 or SEQ ID NO 6;; and

(f) an isolated polypeptide having or comprising a polypeptide sequence that has an Identity Index of 0.95, 0.96, 0.97, 0.98, or 0.99 compared to the polypeptide sequence of SEQ ID NO 2 or SEQ ID NO 6;

(g) fragments and variants of such polypeptides in (a) to (f).

Polypeptides of the present invention are believed to be members of the G protein-coupled receptor family of polypeptides. The biological properties of the respiratory chemokine receptors are hereinafter referred to as "biological activity of respiratory chemokine receptors " or "respiratory chemokine receptor activity" including monocyte/macrophage migration/activation, airway remodeling, airway fibrosis, regulation of epithelial differentiation, regulation of mucus hypersecretion, regulation of mucocilliary clearance,

regulation of inflammation, modulation of neutrophil, T-cell and eosinophil migration and/or activation and regulation of epithelial cell or mast cell activation.
Polypeptides of the present invention also includes variants of the aforementioned polypeptides, including all allelic forms and splice variants. Such polypeptides vary from the reference polypeptide by insertions, deletions, and substitutions that may be conservative or non-conservative, or any combination thereof. Particularly preferred variants are those in which several, for instance from 50 to 30, from 30 to 20, from 20 to 10, from to 5, from 5 to 3, from 3 to 2, from 2 to 1 or 1 amino acids are inserted, substituted, or deleted, in any combination.

Preferred fragments of polypeptides of the present invention include an isolated polypeptide comprising an amino acid sequence having at least 30, 50 or 100 contiguous amino acids from the amino acid sequence of SEQ ID NO 2 or SEQ ID NO. 6 or an isolated polypeptide comprising an amino acid sequence having at least 30, 50 or 100 contiguous amino acids truncated or deleted from the amino acid sequence of SEQ ID NO 2 or SEQ ID NO. 6. Preferred fragments are biologically active fragments that mediate the biological activity of monocyte/macrophage migration/activation, airway remodeling, airway fibrosis, regulation of epithelial differentiation, regulation of mucus hypersecretion, regulation of mucocilliary clearance, regulation of inflammation, modulation of neutrophil, T-cell and eosinophil migration and/or activation and regulation of epithelial cell or mast cell activation, including those with a similar activity or an improved activity, or with a decreased undesirable activity. Also preferred are those fragments that are antigenic or immunogenic in an animal, especially in a human.

Fragments of the polypeptides of the invention may be employed for producing the corresponding full-length polypeptide by peptide synthesis; therefore, these variants may be employed as intermediates for producing the full-length polypeptides of the invention. The polypeptides of the present invention may be in the form of the "mature" protein or maybe a part of a larger protein such as a precursor or a fusion protein. It is often advantageous to include an additional amino acid sequence that contains secretory or leader sequences, pro-sequences, sequences that aid in purification, for instance multiple histidine residues, or an additional sequence for stability during recombinant production.

Polypeptides of the present invention can be prepared in any suitable manner, for instance by isolation form naturally occurring sources, from genetically engineered host cells comprising expression systems (vide infra) or by chemical synthesis, using for instance automated peptide synthesizers, or a combination of such methods. The means for preparing such polypeptides are well understood in the art.

In a further aspect, the present invention relates to respiratory chemokine receptor polynucleotides. Such polynucleotides include:
(a) an isolated polynucleotide comprising a polynucleotide sequence having at least 95%, 96%, 97%, 98%, or 99% identity to the polynucleotide squence of SEQ ID NO. 1 or SEQ ID NO. 5 ; (b) an isolated polynucleotide comprising the polynucleotide of SEQ ID NO. 1 or SEQ ID NO. 5; (c) an isolated polynucleotide having at least 95%, 96%, 97%, 98%, or 99% identity to the polynucleotide of SEQ ID NO 1 or SEQ ID NO. 5 ;
(d) the isolated polynucleotide of SEQ ID NO 1 or SEQ ID NO. 5;
(e) an isolated polynucleotide comprising a polynucleotide sequence encoding a polypeptide sequence having at least 95%, 96%, 97%, 98%, or 99% identity to the polypeptide sequence of SEQ ID NO 2 or SEQ ID NO.6;
(f) an isolated polynucleotide comprising a polynucleotide sequence encoding the polypeptide of SEQ ID NO 2 or SEQ ID NO.6;
(g) an isolated polynucleotide having a polynucleotide sequence encoding a polypeptide sequence having at least 95%, 96%, 97%, 98%, or 99% identity to the polypeptide sequence of SEQ ID NO. 2 or SEQ ID NO.6;
(h) an isolated polynucleotide encoding the polypeptide of SEQ ID NO. 2 or SEQ ID NO.6;
(i) an isolated polynucleotide having or comprising a polynucleotide sequence that has an Identity Index of 0.95, 0.96, 0.97, 0.98, or 0.99 compared to the polynucleotide sequence of SEQ ID NO. 2 or SEQ ID NO.6 an isolated polynucleotide having or comprising a polynucleotide sequence encoding a polypeptide sequence that has an Identity Index of 0.95, 0.96, 0.97, 0.98, or 0.99 compared to the polypeptide sequence of SEQ ID NO 2 or SEQ ID NO.6; and polynucleotides that are fragments and variants of the above mentioned polynucleotides or that are complementary to above mentioned polynucleotides, over the entire length thereof.

Preferred fragments of polynucleotides of the present invention include an isolated polynucleotide comprising an nucleotide sequence having at least 15, 30, 50 or 100

contiguous nucleotides from the sequence of SEQ ID NO 1 or SEQ ID NO 5, or an isolated polynucleotide comprising an sequence having at least 30, 50 or 100 contiguous nucleotides truncated or deleted from the sequence of SEQ ID NO 1 or SEQ ID NO 5.

Preferred variants of polynucleotides of the present invention include splice variants, allelic variants, and polymorphisms,including polynucleotides having one or more single nucleotide polymorphisms (SNPs).

Polynucleotides of the present invention also include polynucleotides encoding polypeptide variants that comprise the amino acid sequence of SEQ ID NO 2 or SEQ ID NO 6 and in which several, for instance from 50 to 30, from 30 to 20, from 20 to 10, from 10 to 5, from 5 to 3, from 3 to 2, from 2 to 1 or 1 amino acid residues are substituted, deleted or added, in any combination.

In a further aspect, the present invention provides polynucleotides that are RNA transcripts of the DNA sequences of the present invention. Accordingly, there is provided an RNA polynucleotide that:
(a) comprises an RNA transcript of the DNA sequence encoding the polypeptide of SEQ ID NO 2 or SEQ ID NO 6;
(b) is the RNA transcript of the, DNA sequence encoding the polypeptide of SEQ ID NO 2 or SEQ ID NO 6;
(c) comprises an RNA transcript of the DNA sequence of SEQ ID NO 1 or SEQ ID NO. 5; or
(d) is the RNA transcript of the DNA sequence of SEQ ID NO 1 or SEQ ID NO 5; and RNA polynucleotides that are complementary thereto.

The polynucleotide sequence of SEQ ID NO 1 is a cDNA sequence that encodes the polypeptide of SEQ ID NO 2. The polynucleotide sequence of SEQ ID NO 5 is a cDNA sequence that encodes the polypeptide of SEQ ID NO 6. The polynucleotide sequence encoding the polypeptide of SEQ ID NO 2 may be identical to the polypeptide encoding sequence of SEQ ID NO 1 or it may- be a sequence other than SEQ ID NO 1, which, as a result of the redundancy (degeneracy) of the genetic code, also encodes the polypeptide of SEQ ID NO 2. The polynucleotide sequence encoding the polypeptide of SEQ ID NO 6 may be identical to the polypeptide encoding sequence of SEQ ID NO 5 or it may- be a sequence other than SEQ ID NO 5, which, as a result of the redundancy (degeneracy) of

the genetic code, also encodes the polypeptide of SEQ ID NO 6. The polypeptide of the SEQ ID NO 2 or SEQ ID NO 6 is related to other proteins of the G protein-coupled receptors family, having homology and/or structural similarity with GPCR - LYMST Jensen, C.P. et al., Proc. Natl. Acad. Sci. U.S.A. 91: 4816-4820,1994).

Preferred polypeptides and polynucleotides of the present invention are expected to have, inter alia, similar biological functions/properties to their homologous polypeptides and polynucleotides. Furthermore, preferred polypeptides and polynucleotides of the present invention have at least one activity such as: monocyte/macrophage migration/activation, airway remodeling, airway fibrosis, regulation of epithelial differentiation, regulation of mucus hypersecretion, regulation of mucocilliary clearance, regulation of inflammation, modulation of neutrophil, T-cell and eosinophil migration and/or activation and regulation of epithelial cell or mast cell activation.

Polynucleotides of the present invention may be obtained using standard cloning and screening techniques from a cDNA library derived from mRNA in cells of human adult or fetal tissue (see for instance, Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989)). Polynucleotides of the invention can also be obtained from natural sources such as genomic DNA libraries or can be synthesized using well known and commercially available techniques.

When polynucleotides of the present invention are used for the recombinant production of polypeptides of the present invention, the polynucleotide may include the coding sequence for the mature polypeptide, by itself, or the coding sequence for the mature polypeptide in reading frame with other coding sequences, such as those encoding a leader or secretory sequence, a pre-, or pro- or prepro- protein sequence, or other fusion peptide portions. For example, a marker sequence that facilitates purification of the fused polypeptide can be encoded. In certain preferred embodiments of this aspect of the invention, the marker sequence is a hexa-histidine peptide, as provided in the pQE vector (Qiagen, Inc.) and described in Gentz et aL, Proc Natl Acad Sci USA (1989) 86:821-824, or is an HA tag. The polynucleotide may also contain non-coding 5' and 3' sequences, such as transcribed, non-translated sequences, splicing and polyadenylation signals, ribosome binding sites and sequences that stabilize mRNA.

Polynucleotides that are identical, or have sufficient identity to a polynucleotide sequence of SEQ ID NO 1 or SEQ ID NO 5 may be used as hybridization probes for cDNA and genomic DNA or as primers for a nucleic acid amplification reaction (for instance, PCR). Such probes and primers may be used to isolate full-length cDNAs and genomic clones encoding polypeptides of the present invention and to isolate cDNA and genomic clones of other genes (including genes encoding paralogs from human sources and orthologs and paralogs from species other than human) that have a high sequence similarity to SEQ ID NO: 1 or SEQ ID NO 5 , typically at least 95% identity. Preferred probes and primers will generally comprise at least 15 nucleotides, preferably, at least 30 nucleotides and may have at least 50, if not at least 100 nucleotides. Particularly preferred probes will have between 30 and 50 nucleotides. Particularly preferred primers will have between 20 and 25 nucleotides.

A polynucleotide encoding a polypeptide of the present invention, including homologs from species other than human, may be obtained by a process comprising the steps of screening a library under stringent hybridization conditions with a labeled probe having the sequence of SEQ ID NO: 1 or SEQ ID NO 5 or a fragment thereof, preferably of at least 15 nucleotides; and isolating full- length cDNA and genomic clones containing said polynucleotide sequence. Such hybridization techniques are well known to the skilled artisan. Preferred stringent hybridization conditions include overnight incubation at 420°C in a solution comprising: 50% formamide, 5xSSC (150mM NaCl, 15mM trisodium citrate), 50 mM sodium phosphate (pH7.6), 5x Denhardt's solution, 10 % dextran sulfate, and 20 microgram/ml denatured, sheared salmon sperm DNA; followed by washing the filters in 0.1x SSC at about 650°C. Thus the present invention also includes isolated polynucleotides, preferably with a nucleotide sequence of at least 100, obtained by screening a library under stringent hybridization conditions with a labeled probe having the sequence of SEQ ID NO: 1 or SEQ ID NO 5 or a fragment thereof, preferably of at least 15 nucleotides.

The person skilled in the art will appreciate that, in many cases, an isolated cDNA sequence will be incomplete, in that the region coding for the polypeptide does not extend all the way through to the 5' terminus. This is a consequence of reverse transcriptase, an enzyme with inherently low processivity (a measure of the ability of the enzyme to remain attached to the

template during the polymerisation reaction), failing to complete a DNA copy of the mRNA template during first strand cDNA synthesis.

There are several methods available and well known to those skilled in the art to obtain full-length cDNAs, or extend short cDNAs, for example those based on the method of Rapid Amplification of cDNA ends (RACE) (see, for example, Frohman et al., Proc Nat Acad Sci USA 85, 8998- 9002, 1988). Recent modifications of the technique, exemplified by the Marathon (trade mark) technology (Clontech Laboratories Inc.) for example, have significantly simplified the search for longer cDNAs. In the Marathon (trade mark) technology, cDNAs have been prepared from mRNA extracted from a chosen tissue and an 'adaptor' sequence ligated onto each end. Nucleic acid amplification (PCR) is then carried out to amplify the "missing" 5' end of the cDNA using a combination of gene specific and adaptor specific oligonucleotide primers. The PCR reaction is then repeated using 'nested' primers, that is, primers designed to anneal within the amplified product (typically an adaptor specific primer that anneals further 3' in the adaptor sequence and a gene specific primer that anneals further 5' in the known gene sequence). The products of this reaction can then be analysed by DNA sequencing and a full-length cDNA constructed either by joining the product directly to the existing cDNA to give a complete sequence, or carrying out a separate full-length PCR using the new sequence information for the design of the 5' primer.

Recombinant polypeptides of the present invention may be prepared by processes well known in the art from genetically engineered host cells comprising expression systems. Accordingly, in a further aspect, the present invention relates to expression systems comprising a polynucleotide or polynucleotides of the present invention, to host cells which are genetically engineered with such expression sytems and to the production of polypeptides of the invention by recombinant techniques. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention.

For recombinant production, host cells can be genetically engineered to incorporate expression systems or portions thereof for polynucleotides of the present invention. Polynucleotides may be introduced into host cells by methods described in many standard

laboratory manuals, such as Davis et al., Basic Methods in Molecular Biology (1986) and Sambrook et al.(ibid).

Preferred methods of introducing polynucleotides into host cells include, for instance, calcium phosphate transfection, DEAE-dextran mediated transfection, microinjection, cationic lipid-mediated transfection, electroporation, transduction, scrape loading, ballistic introduction or infection.

Representative examples of appropriate hosts include bacterial cells, such as Streptococci, Staphylococci, E coli, Streptomyces and Bacillus subtilis cells; fungal cells, such as yeast cells and Aspergillus cells; insect cells such as Drosophila S2 and Spodoptera Sf9 cells; animal cells such as CHO, COS, HeLa, C1 27, 3T3, BHK, HEK 293 and Bowes melanoma cells; and plant cells.

A great variety of expression systems can be used, for instance, chromosomal, episomal and virus-derived systems, e.g., vectors derived from bacterial plasmids, from bacteriophage, from transposons, from yeast episomes, from insertion elements, from yeast chromosomal elements, from viruses such as baculoviruses, papova viruses, such as SV40, vaccinia viruses, adenoviruses, fowl pox viruses, pseudorabies viruses and retroviruses, and vectors derived from combinations thereof, such as those derived from plasmid and bacteriophage genetic elements, such as cosmids and phagemids. The expression systems may contain control regions that regulate as well as engender expression. Generally, any system or vector that is able to maintain, propagate or express a polynucleotide to produce a polypeptide in a host may be used. The appropriate polynucleotide sequence may be inserted into an expression system by any of a variety of well-known and routine techniques, such as, for example, those set forth in Sambrook et al (see above). Appropriate secretion signals may be incorporated into the desired polypeptide to allow secretion of the translated protein into the lumen of the endoplasmic reticulum, the periplasmic space or the extracellular environment. These signals may be endogenous to the polypeptide or they may be heterologous signals.

If a polypeptide of the present invention is to be expressed for use in screening assays, it is generally preferred that the polypeptide be produced at the surface of the cell. In this event, the cells may be harvested prior to use in the screening assay. If the polypeptide is secreted

into the medium, the medium can be recovered in order to recover and purify the polypeptide. If produced intracellularly, the cells must first be lysed before the polypeptide is recovered.

Polypeptides of the present invention can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography is employed for purification. Well known techniques for refolding proteins may be employed to regenerate active conformation when the polypeptide is denatured during intracellular synthesis, isolation and/or purification.

Polynucleotides of the present invention may be used as diagnostic reagents, through detecting mutations in the associated gene. Detection of a mutated form of the gene characterised by the polynucleotide of SEQ ID NO 1 or SEQ ID NO 5 in the cDNA or genomic sequence and which is associated with a dysfunction will provide a diagnostic tool that can add to, or define, a diagnosis of a disease, or susceptibility to a disease, which results from under-expression, over-expression or altered spatial or temporal expression of the gene. Individuals carrying mutations in the gene may be detected at the DNA level by a variety of techniques well known in the art.

Nucleic acids for diagnosis may be obtained from a subject's cells, such as from blood, urine, saliva, tissue biopsy or autopsy material. The genomic DNA may be used directly for detection or it may be amplified enzymatically by using PCR, preferably RT-PCR, or other amplification techniques prior to analysis. RNA or cDNA may also be used in similar fashion. Deletions and insertions can be detected by a change in size of the amplified product in comparison to the normal genotype. Point mutations can be identified by hybridizing amplified DNA to labeled nucleotide sequences. Perfectly matched sequences can be distinguished from mismatched duplexes by RNase digestion or by differences in melting temperatures.

DNA sequence difference may also be detected by alterations in the electrophoretic mobility of DNA fragments in gels, with or without denaturing agents, or by direct DNA sequencing

(see, for instance, Myers et aL, Science (1985) 230:1242). Sequence changes at specific locations may also be revealed by nuclease protection assays, such as RNase and S 1 protection or the chemical cleavage method (see Cotton et aL, Proc Natl Acad Sci USA (1985) 85: 4397-4401).

An array of oligonucleotides probes comprising respiratory chemokine receptor polynucleotide sequence or fragments thereof can be constructed to conduct efficient screening of e.g., genetic mutations. Such arrays are preferably high density arrays or grids. Array technology methods are well known and have general applicability and can be used to address a variety of questions in molecular genetics including gene expression, genetic linkage, and genetic variability, see, for example, M.Chee et al., Science, 274, 610- 613 (1996) and other references cited therein.

Detection of abnormally decreased or increased levels of polypeptide or mRNA expression may also be used for diagnosing or determining susceptibility of a subject to a disease of the invention. Decreased or increased expression can be measured at the RNA level using any of the methods well known in the art for the quantitation of polynucleotides, such as, for example, nucleic acid amplification, for instance PCR, RT- PCR, RNase protection, Northern blotting and other hybridization methods. Assay techniques that can be used to determine levels of a protein, such as a polypeptide of the present invention, in a sample derived from a host are well-known to those skilled in the art. Such assay methods include radioimmunoassays, competitive-binding assays, Western Blot analysis and ELISA assays.

Thus in another aspect, the present invention relates to a diagonostic kit comprising:
(a) a polynucleotide of the present invention, preferably the nucleotide sequence of SEQ ID NO 1 or SEQ ID NO 5, or a fragment or an RNA transcript thereof;
(b) a nucleotide sequence complementary to that of (a);
(c) a polypeptide of the present invention, preferably the polypeptide of SEQ ID NO 2 or SEQ ID NO 6 or a fragment thereof; or
(d) an antibody to a polypeptide of the present invention, preferably to the polypeptide of SEQ ID NO 2 or SEQ ID NO 6.

It will be appreciated that in any such kit, (a), (b), (c) or (d) may comprise a substantial component. Such a kit will be of use in diagnosing a disease or susceptibility to a disease, particularly diseases of the invention, amongst others.

The polynucleotide sequences of the present invention are valuable for chromosome localisation studies. The sequence is specifically targeted to, and can hybridize with, a particular location on an individual human chromosome. The mapping of relevant sequences to chromosomes according to the present invention is an important first step in correlating those sequences with gene associated disease. Once a sequence has been mapped to a precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data. Such data are found in, for example, V. McKusick, Mendelian Inheritance in Man (available on-line through Johns Hopkins University Welch Medical Library). The relationship between genes and diseases that have been mapped to the same chromosomal region are then identified through linkage analysis (co-inheritance of physically adjacent genes). Precise human chromosomal localisations for a genomic sequence (gene fragment etc.) can be determined using Radiation Hybrid (RH) Mapping (Walter, M. Spillett, D., Thomas, P., Weissenbach, J., and Goodfellow, P., (1994) A method for constructing radiation hybrid maps of whole genomes, Nature Genetics 7, 22-28). A number of RH panels are available from Research Genetics (Huntsville, AL, USA) e.g. the - GeneBridge4 RH panel (Hum Moi Genet 1996 Mar;5(3):339-46 A radiation hybrid map of the human genome. Gyapay G, Schmitt K, Fizames C, Jones H, Vega-Czarny N, Spilleft D, Muselet D, Prud'Homme JF, Dib C, Auffray C, Morissette J, Weissenbach J, Goodfellow PN). To determine the chromosomal location of a gene using this panel, 93 PCRs are performed using primers designed from the gene of interest on RH IDNAs. Each of these DNAs contains random human genomic fragments maintained in a hamster background (human / hamster hybrid cell lines). These PCRs result in 93 scores indicating the presence or absence of the PCR product of the gene of interest. These scores are compared with scores created using PCR products from genomic sequences of known location. This comparison is conducted at hftp://www.genome.wi.mit.edu/.
The polynucleotide sequences of the present invention are also valuable tools for tissue expression studies. Such studies allow the determination of expression patterns of polynucleotides of the present invention which may give an indication as to the expression patterns of the encoded polypeptides in tissues, by detecting the mRNAs that encode them. The techniques used are well known in the art and include in situ hydridisation techniques

to clones arrayed on a grid, such as cDNA microarray hybridisation (Schena et al, Science, 270, 467-470, 1995 and Shalon et a], Genome Res, 6, 639-645, 1996) and nucleotide amplification techniques such as PCR. A preferred method uses the TAQMAN (Trade mark) technology available from Perkin Elmer. Results from these studies can provide an indication of the normal function of the polypeptide in the organism. In addition, comparative studies of the normal expression pattern of mRNAs with that of mRNAs encoded by an alternative form of the same gene (for example, one having an alteration in polypeptide coding potential or a regulatory mutation) can provide valuable insights into the role of the polypeptides of the present invention, or that of inappropriate expression thereof in disease. Such inappropriate expression may be of a temporal, spatial or simply quantitative nature.

The polypeptides of the present invention are expressed in respiratory tissues and tissues related to monocyte/macrophage migration/activation, airway remodeling, airway fibrosis, regulation of epithelial differentiation, regulation of mucus hypersecretion, regulation of mucocilliary clearance, regulation of inflammation, modulation of neutrophil, T-cell and eosinophil migration and/or activation and regulation of epithelial cell or mast cell activation.

A further aspect of the present invention relates to antibodies. The polypeptides of the invention or their fragments, or cells expressing them, can be used as immunogens to produce antibodies that are immunospecific for polypeptides of the present invention. The term "immunospecific" means that the antibodies have substantially greater affinity for the polypeptides of the invention than their affinity for other related polypeptides in the prior art. Antibodies generated against polypeptides of the present invention may be obtained by administering the polypeptides or epitope-bearing fragments, or cells to an animal, preferably a non-human animal, using routine protocols. For preparation of monoclonal antibodies, any technique which provides antibodies produced by continuous cell line cultures can be used. Examples include the hybridoma technique (Kohler, G. and Milstein, C., Nature (1975) 256:495-497), the trioma technique, the human B-cell hybridoma technique (Kozbor et aL, Immunology Today (1983) 4:72) and the EBV-hybridoma technique (Cole et aL, Monoclonal Antibodies and Cancer Therapy, 77-96, Alan R. Liss, Inc., 1985).

Techniques for the production of single chain antibodies, such as those described in U.S. Patent No. 4,946,778, can also be adapted to produce single chain antibodies to

polypeptides of this invention. Also, transgenic mice, or other organisms, including other mammals, may be used to express humanized antibodies.

The above-described antibodies may be employed to isolate or to identify clones expressing the polypeptide or to purify the polypeptides by affinity chromatography. Antibodies against polypeptides of the present invention may also be employed to treat diseases of the invention, amongst others.

Polypeptides and polynucleotides of the present invention may also be used as vaccines. Accordingly, in a further aspect, the present invention relates to a method for inducing an immunological response in a mammal that comprises inoculating the mammal with a polypeptide of the present invention, adequate to produce antibody and/or T cell immune response, including, for example, cytokine-producing T cells or cytotoxic T cells, to protect said animal from disease, whether that disease is already established within the individual or not. An immunological response in a mammal may also be induced by a method comprises delivering a polypeptide of the present invention via a vector directing expression of the poiynucleotide and coding for the polypeptide in vivo in order to induce such an immunological response to produce antibody to protect said animal from diseases of the invention. One way of administering the vector is by accelerating it into the desired cells as a coating on particles or otherwise. Such nucleic acid vector may comprise DNA, RNA, a modified nucleic acid, or a DNA/RNA hybrid. For use a vaccine, a polypeptide or a nucleic acid vector will be normally provided as a vaccine formulation (composition). The formulation may further comprise a suitable carrier. Since a polypeptide may be broken down -in the stomach, it is preferably administered parenterally (for instance, subcutaneous, intramuscular, intravenous, or intradermal injection). Formulations suitable for parenteral administration include aqueous and non-aqueous sterile injection solutions that may contain anti-oxidants, buffers, bacteriostats and solutes that render the formulation instonic with the blood of the recipient; and aqueous and non-aqueous sterile suspensions that may include suspending agents or thickening agents.

The formulations may be presented in unit-dose or multi-dose containers, for example, sealed ampoules and vials and may be stored in a freeze- dried condition requiring only the addition of the sterile liquid carrier immediately prior to use. The vaccine formulation may also include adjuvant systems for enhancing the immunogenicity of the formulation, such as

oil-in water systems and other systems known in the art. The dosage will depend on the specific activity of the vaccine and can be readily determined by routine experimentation.

Polypeptides of the present invention have one or more biological functions that are of relevance in one or more disease states, in particular the diseases of the invention hereinbefore mentioned. It is therefore useful to identify compounds that stimulate or inhibit the function or level of the polypeptide. Accordingly, in a further aspect, the present invention provides for a method of screening compounds to identify those that stimulate or inhibit the function or level of the polypeptide. Such methods identify agonists or antagonists that may be employed for therapeutic and prophylactic purposes for such diseases of the invention as hereinbefore mentioned. Compounds may be identified from a variety of sources, for example, cells, cell-free preparations, chemical libraries, collections of chemical compounds, and natural product mixtures. Such agonists or antagonists so-identified may be natural or modified substrates, ligands, receptors, enzymes, etc., as the case may be, of the polypeptide; a structural or functional mimetic thereof (see Coligan et aL, Current Protocols in Immunology 1(2):Chapter 5 (1991)) or a small molecule.

The screening method may simply measure the binding of a candidate compound to the polypeptide, or to cells or membranes bearing the polypeptide, or a fusion protein thereof, by means of a label directly or indirectly associated with the candidate compound. Alternatively, the screening method may involve measuring or detecting (qualitatively or quantitatively) the competitive binding of a candidate compound to the polypeptide against a labeled competitor (e.g. agonist or antagonist). Further, these screening methods may test whether the candidate compound results in a signal generated by activation or inhibition of the polypeptide, using detection systems appropriate to the cells bearing the polypeptide. Inhibitors of activation are generally assayed in the presence of a known agonist and the effect on activation by the agonist by the presence of the candidate compound is observed. Further, the screening methods may simply comprise the steps of mixing a candidate compound with a solution containing a polypeptide of the present invention, to form a mixture, measuring a HGRL101 activity in the mixture, and comparing the HGRL101 activity of the mixture to a control mixture which contains no candidate compound.

Polypeptides of the present invention may be employed in conventional low capacity screening methods and also in high-throughput screening (HTS) formats. Such HTS formats include not only the well-established use of 96- and, more recently, 384-well micotiter plates but also emerging methods such as the nanowell method described by Schullek et al, Anal Biochem., 246, 20-29, (1997).

Fusion proteins, such as those made from Fc portion and respiratory chemokine receptor polypeptide, as hereinbefore described, can also be used for high-throughput screening assays to identify antagonists for the polypeptide of the present invention (see D. Bennett et aL, J Mol Recognition, 8:52-58 (1995); and K. Johanson et al., J Biol Chem, 270(16):9459-9471 (1995)).

Screening techniques

The polynucleotides, polypeptides and antibodies to the polypeptide of the present invention may also be used to configure screening methods for detecting the effect of added compounds on the production of mRNA and polypeptide in cells. For example, an ELISA assay may be constructed for measuring secreted or cell associated levels of polypeptide using monoclonal and polyclonal antibodies by standard methods known in the art. This can be used to discover agents that may inhibit or enhance the production of polypeptide (also called antagonist or agonist, respectively) from suitably manipulated cells or tissues.

A polypeptide of the present invention may be used to identify membrane bound or soluble receptors, if any, through standard receptor binding techniques known in the art. These include, but are not limited to, ligand binding and crosslinking assays in which the polypeptide is labeled with a radioactive isotope (for instance, 1251), chemically modified (for instance, biotinylated), or fused to a peptide sequence suitable for detection or purification, and incubated with a source of the putative receptor (cells, cell membranes, cell supernatants, tissue extracts, bodily fluids). Other methods include biophysical techniques such as surface plasmon resonance and spectroscopy. These screening methods may also be used to identify agonists and antagonists of the polypeptide that compete with the binding of the polypeptide to its receptors, if any. Standard methods for conducting such assays are well understood in the art.

Examples of antagonists of polypeptides of the present invention include antibodies or, in some cases, oligonucleotides or proteins that are closely related to the ligands, substrates, receptors, enzymes, etc., as the case may be, of the poiypeptide, e.g., a fragment of the ligands, substrates, receptors, enzymes, etc.; or a small molecule that bind to the polypeptide of the present invention but do not elicit a response, so that the activity of the polypeptide is prevented.

Screening methods may also involve the use of transgenic technology and a respiratory chemokine receptor gene. The art of constructing transgenic animals is well established. For example, the respiratory chemokine receptor gene may be introduced through microinjection into the male pronucleus of fertilized oocytes, retroviral transfer into pre- or post-implantation embryos, or injection of genetically modified, such as by electroporation, embryonic stem cells into host blastocysts. Particularly useful transgenic animals are so-called "knock-in" animals in which an animal gene is replaced by the human equivalent within the genome of that animal. Knock-in transgenic animals are useful in the drug discovery process, for target validation, where the compound is specific for the human target. Other useful transgenic animals are so-called "knock-out" animals in which the expression of the animal ortholog of a polypeptide of the present invention and encoded by an endogenous DNA sequence in a cell is partially or completely annulled. The gene knock-out may be targeted to specific cells or tissues, may occur only in certain cells or tissues as a consequence of the limitations of the technology, or may occur in all, or substantially all, cells in the animal. Transgenic animal technology also offers a whole animal expression-cloning system in which introduced genes are expressed to give large amounts of polypeptides of the present invention.

Screening kits for use in the above described methods form a further aspect of the present invention. Such screening kits comprise:
(a) a polypeptide of the present invention;
(b) a recombinant cell expressing a polypeptide of the present invention,
(c) a cell membrane expressing a polypeptide of the present invention; or
(d) an antibody to a polypeptide of the present invention; which polypeptide is preferably that of SEQ ID NO 2 or SEQ ID. No.6.
It will be appreciated that in any such kit, (a), (b), (c) or (d) may comprise a substantial component.

Glossary

The following definitions are provided to facilitate understanding of certain terms used frequently hereinbefore.

"Antibodies" as used herein includes polyclonal and monoclonal antibodies, chimeric, single chain, and humanized antibodies, as well as Fab fragments, including the products of an Fab or other immunoglobulin expression library.

"Isolated" means altered by the human hands from its natural state, ie. if it occurs in nature, it has been changed or removed from its original environment, or both. For example, a polynucleotide or a polypeptide naturally present in a living organism is not "isolated," but the same polynucleotide or polypeptide separated from the coexisting materials of its natural state is "isolated", as the term is employed herein. Moreover, a polynucleotide or polypeptide that is introduced into an organism by transformation, genetic manipulation or by any other recombinant method is "isolated" even if it is still present in said organism, which organism may be living or non-living.

"Polynucleotide" generally refers to any polyribonucleotide (RNA) or polydeoxribonucleotide (DNA), which may be unmodified or modified RNA or DNA. "Polynucleotides" include, without limitation, single- and double-stranded DNA, DNA that is a mixture of single- and double- stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded or a mixture of single- and double-stranded regions. In addition, "polynucleotide" refers to triple-stranded regions comprising RNA or DNA or both RNA and DNA. The term "polynucleotide" also includes DNAs or RNAs containing one or more modified bases and DNAs or RNAs with backbones modified for stability or for other reasons.

"Modified" bases include, for example, tritylated bases and unusual bases such as inosine. A variety of modifications may be made to DNA and RNA; thus, "polynucleotide" embraces

chemically, enzymatically or metabolically modified forms of polynucleotides as typically found in nature, as well as the chemical forms of DNA and RNA characteristic of viruses and cells. "Polynucleotide" also embraces relatively short polynucleotides, often referred to as oligonucleotides.

"Polypeptide" refers to any polypeptide comprising two or more amino acids joined to each other by peptide bonds or modified peptide bonds, - i.e., peptide isosteres. "Polypeptide" refers to both short chains, commonly referred to as peptides, oligopeptides or oligomers, and to longer chains, generally referred to as proteins. Polypeptides may contain amino acids other than the 20 gene-encoded amino acids. "Polypeptides" include amino acid sequences modified either by natural processes, such as post-translational processing, or by chemical modification techniques that are well known in the art. Such modifications are well described in basic texts and in more detailed monographs, as well as in a voluminous research literature. Modifications may occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini.
It will be appreciated that the same type of modification may be present to the same or varying degrees at several sites in a given polypeptide. Also, a given polypeptide may contain many types of modifications. Polypeptides may be branched as a result of ubiquitination, and they may be cyclic, with or without branching. Cyclic, branched and branched cyclic polypeptides may result from post-translation natural processes or may be made by synthetic methods. Modifications include acetylation, acylation, ADP-ribosylation, amidation, biotinylation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphotidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cystine, formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination (see, for instance, Proteins - Structure and Molecular Properties, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York, 1993; Wold, F., Post-translational Protein Modifications: Perspectives and Prospects, 1-12, in Post-translational Covalent Modification of Proteins, B. C. Johnson, Ed., Academic Press, New York, 1983; Seifter et aL, "Analysis for protein modifications and nonprotein cofactors", Meth Enzymol,

182, 626-646, 1990, and Rattan et al., "Protein Synthesis: Post-translational Modifications and Aging", Ann NY Acad Sci, 663, 48-62, 1992).

"Fragment" of a polypeptide sequence refers to a polypeptide sequence that is shorter than the reference sequence but that retains essentially the same biological function or activity as the reference polypeptide. "Fragment" of a polynucleotide sequence refers to a polynucloetide sequence that is shorter than the reference sequence of SEQ ID NO 1 or SEQ ID NO 5.

"Variant" refers to a polynucleotide or polypeptide that differs from a reference polynucleotide or polypeptide, but retains the essential properties thereof. A typical variant of a polynucleotide differs in nucleotide sequence from the reference polynucleotide. Changes in the nucleotide sequence of the variant may or may not alter the amino acid sequence of a polypeptide encoded by the reference polynucleotide. Nucleotide changes may result in amino acid substitutions, additions, deletions, fusions and truncations in the polypeptide encoded by the reference sequence, as discussed below. A typical variant of a polypeptide differs in amino acid sequence from the reference polypeptide. Generally, alterations are limited so that the sequences of the reference polypeptide and the variant are closely similar overall and, in many regions, identical. A variant and reference polypeptide may differ in amino acid sequence by one or more substitutions, insertions, deletions in any combination. A substituted or inserted amino acid residue may or may not be one encoded by the genetic code. Typical conservative substitutions include Gly, Ala; Val, lie, Leu; Asp, Glu; Asn, Gln-I Ser, Thr; Lys, Arg; and Phe and Tyr. A variant of a polynucleotide or polypeptide may be naturally occurring such as an allele, or it may be a variant that is not known to occur naturally. Non-naturally occurring variants of polynucleotides and polypeptides may be made by mutagenesis techniques or by direct synthesis. Also included as variants are polypeptides having one or more post-translational modifications, for instance glycosylation, phosphorylation, methylation, ADIP ribosylation and the like. Embodiments include methylation of the N-terminal amino acid, phosphorylations of serines and threonines and modification of C- terminal glycines.

"Allele" refers to one of two or more alternative forms of a gene occuring at a given locus in the genome.

"Polymorphism" refers to a variation in nucleotide sequence (and encoded polypeptide sequence, if relevant) at a given position in the genome within a population.

"Single Nucleotide Polymorphism" (SNP) refers to the occurence of nucleotide variability at a single nucleotide position in the genome, within a population. An SNP may occur within a gene or within intergenic regions of the genome. SNPs can be assayed using Allele Specific Amplification (ASA). For the process at least 3 primers are required. A common primer is used in reverse complement to the polymorphism being assayed. This common primer can be between 50 and 1500 bps from the polymorphic base. The other two (or more) primers are identical to each other except that the final 3' base wobbles to match one of the two (or more) alleles that make up the polymorphism. Two (or more) PCR reactions are then conducted on sample DNA, each using the common primer and one of the Allele Specific Primers.

"Splice Variant" as used herein refers to cDNA molecules produced from RNA molecules initially transcribed from the same genomic DNA sequence but which have undergone alternative RNA splicing. Alternative RNA splicing occurs when a primary RNA transcript undergoes splicing, generally for the removal of introns, which results in the production of more than one mRNA molecule each of that may encode different amino acid sequences. The term splice variant also refers to the proteins encoded by the above cDNA molecules.

"Identity" reflects a relationship between two or more polypeptide sequences or two or more polynucleotide sequences, determined by comparing the sequences. In general, identity refers to an exact nucleotide to nucleotide or amino acid to amino acid correspondence of the two polynucleotide or two polypeptide sequences, respectively, over the length of the sequences being compared.

"% Identity" - For sequences where there is not an exact correspondence, a "% identity" may be determined. In general, the two sequences to be compared are aligned to give a maximum correlation between the sequences. This may include inserting "gaps" in either one or both sequences, to enhance the degree of alignment. A % identity may be determined over the whole length of each of the sequences being - 23 compared (so-called global alignment), that is particularly suitable for sequences of the same or very similar

length, or over shorter, defined lengths (so-called local alignment), that is more suitable for sequences of unequal length.

"Similarity" is a further, more sophisticated measure of the relationship between two polypeptide sequences. In general, "similarity" means a comparison between the amino acids of two polypeptide chains, on a residue by residue basis, taking into account not only exact correspondences between a between pairs of residues, one from each of the sequences being compared (as for identity) but also, where there is not an exact correspondence, whether, on an evolutionary basis, one residue is a likely substitute for the other. This likelihood has an associated "score" from which the "% similarity" of the two sequences can then be determined.

Methods for comparing the identity and similarity of two or more sequences are well known in the art. Thus for instance, programs available in the Wisconsin Sequence Analysis Package, version 9.1 (Devereux J et al, Nucleic Acids Res, 12, 387-395, 1984, available from Genetics Computer Group, Madison, Wisconsin, USA), for example the programs BESTFIT and GAP, may be used to determine the % identity between two polynucleotides and the % identity and the % similarity between two polypeptide sequences. BESTFIT uses the "local homology" algorithm of Smith and Waterman (J Mol Biol, 147,195-197, 1981, Advances in Applied Mathematics, 2, 482-489, 1981) and finds the best single region of similarity between two sequences. BESTFIT is more suited to comparing two polynucleotide or two polypeptide sequences that are dissimilar in length, the program assuming that the shorter sequence represents a portion of the longer. In comparison, GAP aligns two sequences, finding a "maximum similarity", according to the algorithm of Neddleman and Wunsch (J Mol Biol, 48, 443-453, 1970). GAP is more suited to comparing sequences that are approximately the same length and an alignment is expected over the entire length. Preferably, the parameters "Gap Weight" and "Length Weight" used in each program are 50 and 3, for polynucleotide sequences and 12 and 4 for polypeptide sequences, respectively. Preferably, % identities and similarities are determined when the two sequences being compared are optimally aligned.

Other programs for determining identity and/or similarity between sequences are also known in the art, for instance the BLAST family of programs (Altschul S F et al, J Mol Biol, 215, 403-410, 1990, Altschul S F et al, Nucleic Acids Res., 25:389-3402, 1997, available

from the National Center for Biotechnology Information (NCBI), Bethesda, Maryland, USA and accessible through the home page of the NCBI at www.ncbi.nlm.nih.gov) and FASTA (Pearson W R, Methods in Enzymology, 183, 63-99, 1990; Pearson W R and Lipman D J, Proc Nat Acad Sci USA, 85, 2444-2448,1988, available as part of the Wisconsin Sequence Analysis Package).

Preferably, the BLOSUM62 amino acid substitution matrix (Henikoff S and Henikoff J G, Proc. Nat. Acad Sci. USA, 89, 10915-10919, 1992) is used in polypeptide sequence comparisons including where nucleotide sequences are first translated into amino acid sequences before comparison.

Preferably, the program BESTFIT is used to determine the % identity of a query polynucleotide or a polypeptide sequence with respect to a reference polynucleotide or a polypeptide sequence, the query and the reference sequence being optimally aligned and the parameters of the program set at the default value, as hereinbefore described.

"Identity Index" is a measure of sequence relatedness which may be used to compare a candidate sequence (polynucleotide or polypeptide) and a reference sequence. Thus, for instance, a candidate polynucleotide sequence having, for example, an Identity Index of 0.95 compared to a reference polynucleotide sequence is identical to the reference sequence except that the candidate polynucleotide sequence may include on average up to five differences per each 100 nucleotides of the reference sequence. Such differences are selected from the group consisting of at least one nucleotide deletion, substitution, including transition and transversion, or insertion. These differences may occur at the 5' or 3' terminal positions of the reference polynucleotide sequence or anywhere between these terminal positions, interspersed either individually among the nucleotides in the reference sequence or in one or more contiguous groups within the reference sequence. In other words, to obtain a polynucleotide sequence having an Identity Index of 0.95 compared to a reference polynucleotide sequence, an average of up to 5 - 25 in every 100 of the nucleotides of the in the reference sequence may be deleted, substituted or inserted, or any combination thereof, as hereinbefore described. The same applies mutatis mutandis for other values of the Identity Index, for instance 0.96, 0.97, 0.98 and 0.99.

Similarly, for a polypeptide, a candidate polypeptide sequence having, for example, an Identity Index of 0.95 compared to a reference polypeptide sequence is identical to the reference sequence except that the polypeptide sequence may include an average of up to five differences per each 100 amino acids of the reference sequence. Such differences are selected from the group consisting of at least one amino acid deletion, substitution, including conservative and non-conservative substitution, or insertion. These differences may occur at the amino- or carboxy-terminal positions of the reference polypeptide sequence or anywhere between these terminal positions, interspersed either individually among the amino acids in the reference sequence or in one or more contiguous groups within the reference sequence. In other words, to obtain a polypeptide sequence having an Identity Index of 0.95 compared to a reference polypeptide sequence, an average of up to 5 in every 100 of the amino acids in the reference sequence may be deleted, substituted or inserted, or any combination thereof, as hereinbefore described. The same applies mutatis mutandis for other values of the Identity Index, for instance 0.96, 0.97, 0.98 and 0.99. The relationship between the number of nucleotide or amino acid differences and the

Identity Index may be expressed in the following equation:

$$n_a \leq x_a - (x_a \bullet I)$$

in which:

n_a is the number of nucleotide or amino acid differences,

x_a is the total number of nucleotides in SEQ ID NO 1 or SEQ ID No.5, or amino acids in SEQ ID NO.2 or SEQ ID NO 6, respectively,

I is the Identity Index,

• is the symbol for the multiplication operator, and in which any non-integer product Of x_a and I is rounded down to the nearest integer prior to subtracting it from x_a.

"Homolog" is a generic term used in the art to indicate a polynucleotide or polypeptide sequence possessing a high degree of sequence relatedness to a reference sequence. Such relatedness may be quantified by determining the degree of identity and/or similarity between the two sequences as hereinbefore defined. Falling within this generic term are the terms "ortholog", and "paralog". "Ortholog" refers to a polynucleotide or polypeptide that is the functional equivalent of the polynucleotide or polypeptide in another species. "Paralog"

refers to a polynucleotideor polypeptide that within the same species which is functionally similar.

"Fusion protein" refers to a protein encoded by two, unrelated, fused genes or fragments thereof. Examples have been disclosed in US 5541087, 5726044. In the case of Fc-PGPCR-3, employing an immunoglobulin Fc region as a part of a fusion protein is advantageous for performing the functional expression of Fc-PGPCR-3 or fragments of PGPCR-3, to improve pharmacokinetic properties of such a fusion protein when used for therapy and to generate a dimeric Fc-PGPCR-3. The Fc- PGPCR-3 DNA construct comprises in 5' to 3' direction, a secretion cassette, i.e. a signal sequence that triggers export from a mammalian cell, DNA encoding an immunoglobulin Fc region fragment, as a fusion partner, and a DNA encoding Fc-PGPCR-3 or fragments thereof. In some uses it would be desirable to be able to alter the intrinsic functional properties (complement binding, Fc-Receptor binding) by mutating the functional Fc sides while leaving the rest of the fusion protein untouched or delete the Fc part completely after expression.

All publications and references, including but not limited to patents and patent applications, cited in this specification are herein incorporated by reference in their entirety as if each individual publication or reference were specifically and individually indicated to be incorporated by reference herein as being fully set forth. Any patent application to which this application claims priority is also incorporated by reference herein in its entirety in the manner described above for publications and references.

Examples

Example 1

Mammalian cell expression

The receptors of the present invention are expressed in either human embryonic kidney 293 (HEK293) cells or adherent dhfr CHO cells. To maximize receptor expression, typically all 5' and 3' untranslated regions (UTRs) are removed from the receptor cDNA prior to insertion into a pCDN or pCDNA3 vector. The cells are transfected with individual receptor cDNAs by lipofectin and selected in the presence of 400 µg/ml G418. After 3 weeks of selection, individual clones are picked and expanded for further analysis. HEK293 or CHO cells transfected with the vector alone serve as negative controls. To isolate cell lines stably

expressing the individual receptors, about 24 clones are typically selected and analyzed by Northern blot analysis. Receptor mRNAs are generally detectable in about 50% of the G418-resistant clones analyzed.

Example 2

Ligand bank for binding and functional assays.

A bank of over 1200 putative receptor ligands has been assembled for screening. The bank comprises: transmitters, hormones and chemokines known to act via a human seven transmembrane (7TM) receptor; naturally occurring compounds which may be putative agonists for a human 7TM receptor, non-mammalian, biologically active peptides for which a mammalian counterpart has not yet been identified; and compounds not found in nature, but which activate 7TM receptors with unknown natural ligands. This bank is used to initially screen the receptor for known ligands, using both functional (i.e. calcium, cAMP, microphysiometer, oocyte electrophysiology, etc, see below) as well as binding assays.

Example 3

Ligand binding assays

Ligand binding assays provide a direct method for ascertaining receptor pharmacology and are adaptable to a high throughput format. The purified ligand for a receptor is radiolabeled to high specific activity (50-2000 Ci/mmol) for binding studies. A determination is then made that the process of radiolabeling does not diminish the activity of the ligand towards its receptor. Assay conditions for buffers, ions, pH and other modulators such as nucleotides are optimized to establish a workable signal to noise ratio for both membrane and whole cell receptor sources. For these assays, specific receptor binding is defined as total associated radioactivity minus the radioactivity measured in the presence of an excess of unlabeled competing ligand. Where possible, more than one competing ligand is used to define residual nonspecific binding.

Example 4

Functional assay in Xenopus oocytes

Capped RNA transcripts from linearized plasmid templates encoding the receptor cDNAs of the invention are synthesized in vitro with RNA poiymerases in accordance with standard procedures. In vitro transcripts are suspended in water at a final concentration of 0.2 mg/ml.

Ovarian lobes are removed from adult female toads, Stage V defolliculated oocytes are obtained, and RNA transcripts (10 ng/oocyte) are injected in a 50 nl bolus using a microinjection apparatus. Two electrode voltage clamps are used to measure the currents from individual Xenopus oocytes in response to agonist exposure. Recordings are made in Ca2+ free Barth's medium at room temperature. The Xenopus system can be used to screen known ligands and tissue/cell extracts for activating ligands.

Example 5

Microphysiometric assays

Activation of a wide variety of secondary messenger systems results in extrusion of small amounts of acid from a cell. The acid formed is largely as a result of the increased metabolic activity required to fuel the intracellular signaling process. The pH changes in the media surrounding the cell are very small but are detectable by the CYTOSENSOR microphysiometer (Molecular Devices Ltd., Menlo Park, CA). The CYTOSENSOR is thus capable of detecting the activation of a receptor which is coupled to an energy utilizing intracellular signaling pathway such as the G protein-coupled receptor of the present invention.

Example 6

Extract/cell supernatant screening

A large number of mammalian receptors exist for which there remains, as yet, no cognate activating ligand (agonist). Thus, active ligands for these receptors may not be included within the ligands banks as identified to date.

Accordingly, the 7TM receptor of the invention is also functionally screened (using calcium, cAMP, microphysiometer, oocyte electrophysiology, etc., functional screens) against tissue extracts to identify natural ligands.

Extracts that produce positive functional responses can be sequencially subfractionated until an activating ligand is isolated identified.

Example 7

Calcium and cAMP functional assays 7TM receptors which are expressed in HEK 293 cells have been shown to be coupled functionally to activation of PLC and calcium mobilization and/or cAMP stirnuation or inhibition. Basal calcium levels in the HEK 293 cells in receptor-

transfected or vector control cells were observed to be in the normal, 100 nM to 200 nM, range. HEK 293 cells expressing recombinant receptors are loaded with fura 2 and in a single day > 150 selected ligands or tissue/cell extracts are evaluated for agonist induced calcium mobilization. Similarly, HEK 293 cells expressing recombinant receptors are evaluated for the stimulation or inhibition of cAMP production using standard cAMP quantitation assays. Agonists presenting a calcium transient or cAMP flucuation are tested in vector control cells to determine if the response is unique to the transfected cells expressing receptor.

Example 8-Tissue Expression

Primer design

Primers are designed to orphan GPCR sequences either using the PrimerExpress programme (Applied Biosystems) or manually. Primer annealing temperatures and amplification conditions are optimised using genomic DNA. Each reaction mixture contains 0.2 mM dNTP's, 1x PCR buffer containing 1.5 mM $MgCl_2$ (Roche Molecular Biochemicals), 0.5 Units Taq DNA polymerase (Roche Molecular Biochemicals), 223 ng genomic DNA (Promega), 50 pmol each primer and deionised water in a total volume of 25 µl. Cycling is carried out in 0.2 ml tubes using a Biometra T gradient PCR machine with typically 8 reactions covering the annealing temperature range 48ºC to 72.2ºC. Cycling conditions were as follows: Denaturation at 94ºC for 1 min 45s for 1 cycle then 35 cycles of denaturation at 94ºC for 15s, annealing for 15s, extension at 72ºC for 30s. Reactions are analysed on a 1.5% agarose gel.

Preparation of cells, RNA and first strand cDNA synthesis

RNA is prepared from cells using RNeasy extraction kits (Qiagen) according to the manufacturers protocol. Cells used are primary human bronchial epithelial cells (HBEC's), differentiated HBEC's, primary human lung fibroblasts, primary human bronchial smooth muscle cells (SMC), human peripheral blood T-cells, human peripheral blood neutrophils and neutrophils which had been stimulated with GM-CSF. Primary human cells are either purchased from Biowhittaker or, where indicated, isolated from peripheral human blood according to standard procedures. Human macrophages are prepared from *in vitro* differentiated human peripheral blood monocytes according to standard protocols.

First strand cDNA is prepared from total RNA isolated from cells using the reagents and protocol provided in the first strand cDNA synthesis kit (Roche Molecular Biochemicals).

RT-PCR

Selected sequences are profiled in cDNA derived from tissues and in the different cell types described above by reverse transcriptase polymerase chain reaction (RT-PCR). Tissue cDNA's used for RT-PCR profiling are purchased from Clontech. Each reaction mixture contains 0.2 mM dNTP's, 1x PCR buffer containing 1.5 mM $MgCl_2$, 0.5 Units Taq DNA polymerase, 50 pmol each primer and deionised water in a total volume of 25 µl. Template cDNA used is either from commercial cDNA derived from tissue samples from Clontech panel I and II (2.5 µl) or from cDNA's prepared from cell types as described in the previous section (1 µl). Cycling is carried out in 0.2 ml tubes using a Biometra Trio PCR machine with the optimally determined annealing temperature. Cycling conditions are as follows: Denaturation at 94ºC for 1 min 45s for 1 cycle then 35 cycles of denaturation at 94ºC for 15s, annealing for 15s, extension at 72ºC for 30s. Reactions are analysed on a 1.5% agarose gel and stained with ethidium bromide. Control RT-PCR reactions are performed with primers specific to the housekeeping gene GAPDH.

The primer sequences for polynucleotide 413 are SEQ ID NO.3 (for the forward) and SEQ ID NO 4 (for the reverse) and for polynucleotide 88.1 are SEQ ID NO.7 (for the forward) and SEQ ID NO 8 (for the reverse). The tissue distribution is given in Table 1.

Sequencing PCR products

PCR products are excised from agarose gels and purified using the QIAquick spin gel extraction kit (QIAGEN) according to the manufacturer's instructions. PCR products are eluted in 30 µl of sterile water. Typically 10 ng of purified PCR product is sequenced using one of the primers used for amplification and using the BigDye terminator cycle sequencing mix (Applied Biosystems) according to the manufacturer's instructions. Sequencing reactions are analysed on an ABI310 sequencer.

Northern Blot analysis

Northern blot analyses are performed using 12-lane multiple tissue northern blots (Clontech) according to the manufacturer's instructions. Briefly, gel purified cDNA fragments (25 ng) were labelled with fresh [α-^{32}P]dCTP using the Rediprime labelling kit (Amersham)

according to the manufacturer's instructions. Labelled probes are denatured at 95ºC for 5 min then cooled on ice for 5 min immediately prior to use. Prehybridisation is carried out in 5 ml ExpressHyb solution (Clontech) containing 0.1 mg/ml denatured herring sperm DNA at 68ºC for 30 min in a Hybaid oven. Hybridisation is carried out with fresh ExpressHyb solution containing herring sperm DNA (0.1 mg/ml) and denatured probe. Hybridisation is performed at 68°C for approximately 2 h. Blots are washed four times with 2xSSC, 0.05% SDS at RT for 10 min followed by 2 washes in 0.1xSSC, 0.1% SDS for 20 min at 50°C. Blots are sealed in plastic bags and exposed to a phosphoimager screen (Molecular Dynamics) overnight. Images are processed using a Storm Phosphoimager machine (Molecular Dynamics) and with ImageQuant 5.0 software. Blots are stripped by boiling for 10 min in 0.5% SDS and re-hybridised with a β-actin (Clontech) ^{32}P-labelled probe.

Tissue distribution

PCR Tissue Expression Profiles	Code 88.1	Code 413
brain	-	-
heart	-	-
kidney	-	++
liver	-	+
lung	-	-
pancreas	-	-
placenta	-	-
skeletal muscle	-	-
colon	-	-
ovary	-	+
leukocyte	-	-
prostate	-	-
small intestine	-	-
spleen	-	-
testis	-	-
thymus	-	-
human bronchial epithelial cells	+	-
differentiated human bronchial epithelial cells	+++	+
fibroblasts	+	-

bronchial smooth muscle cells	+	-
mixed t-cells	-	+
macrophages	-	+++
neutrophils	-	+++

Tissue Expression

- \- non detected
- \+ weak signal
- ++ good signal
- +++ strong signal

Example 9: BLAST Homology Search

The percentage homology at the nucleotide and amino acid level compared to identified or putative genes:

ORP	Top match Genbank NR (nt)	% ID	Top match Genbank NR (aa)	% ID
Singlet_413	-		Putative chemokine receptor (XP_015921.1) HM74 (D10923)	32 32
Contig_88:1	Putative chemokine receptor FKSG80 (AF345568)	100	Putative chemokine receptor (AF345568) HM74 (D10923)	96 50

Example 10- Singlet_413 Receptor stably expressed in a CHO-K1-CRE cell line

Ligand fishing consists of looking for compounds able to produce a transient Ca^{2+} signal in CHO-K1-CRE cells stably transfected with the orphan receptor singlet_413. Using the $Flipr^{384}$, the compounds are injected on the cells and their signals (kinetics) are simultaneously recorded. Twenty-four hours before the experiment, the cells are havested, counted and seeded in 384-well plates at 10000 cells per wells and put at 37°C/5% CO_2. The day of the experiment, the injection plates containing the compounds are prepared as described in the following table:

ORP-INJECTION-PLATE (384-well)	Quadrant	Plate Name (96-well)
Injection-1	Q1	Peptide_1
	Q2	Peptide_2
	Q3	Peptide_3
	Q4	Peptide_4
Injection-2	Q1	Peptide_5
	Q2	Lipids_1
	Q3	Lipids_2
	Q4	Lipids_3
Injection-3	Q1	CABA 00004925
	Q2	CABA 00005027
	Q3	CABA 00005038
	Q4	CABA 00005039
Injection-4	Q1	CABA 00005040
	Q2	CABA 00005261
	Q3	CABA 00005267
	Q4	CABA 00005269

The cells are loaded with 2 µM Fluo-4 in HBSS/Hepes buffer. After one hour incubation at 37°C/5%CO_2, the cells are washed with HBSS/Hepes containing 2.5 mM probenecid (2 wash cycles + volume adjustment are performed). Plates that will be primed with ATP are adjusted to 20 µl final volume and plates that will directly receive the compounds are adjusted to 40 µl final volume. Four assay plates are primed with 20 µl ATP at 20 µM initial (10 µM final concentration) and are incubated at room temperature for 30 minutes. Compounds plates at 3 times the final concentration are then injected with the Flipr pipettor on both primed and not primed plates. The Flipr laser is setup on 0.6 watt, the exposure length at 0.4 sec. The excitation wavelengh is 488 nM. A 10 seconds baseline is recorded prior compounds injection. After the injection, the signal is recorded every second for one minute. Then, signal decay is measured for 20 additional seconds. Two statistics are exported: Fbasal and Fmax. For each well, a dF/F value is expressed with the following formula : (Fmax-Fbasal)/Fbasal where Fbasal is the basal fluorescence prior compound injection and Fmax is the peak value of the kinetic.

CHO-CRE-Singlet_413

w/o priming : no hit
with priming : 15 hits

np = not primed
p = primed

				CHO-CRE							
				CHO-K1-ATCC		CHO-K1		Contig 88:1		Singlet_413	
Cpds #	Plate	96-well	Name	np	p	np	p	np	p	np	p
1	Lip1	B01	5(S)-HETE	0.02	-0.03	-0.02	0.08	0.00	-0.01	0.01	1.06
2	Lip1	B02	(±)5-HETE	0.02	-0.02	-0.05	-0.05	0.01	-0.01	0.01	1.05
3	Lip1	B06	11(S)-HETE	-0.04	-0.05	-0.03	-0.03	0.00	-0.01	0.00	0.85
4	Lip1	B11	(±)5-HETrE	0.01	-0.04	-0.01	-0.04	0.01	0.02	0.03	0.83
5	Lip1	C02	(±)5-HEPE	0.03	-0.04	-0.04	-0.05	0.01	0.01	0.03	1.03
6	Lip1	C04	5(S)-HPETE	0.03	-0.03	-0.01	0.08	0.00	0.01	0.01	1.30
7	Lip1	C11	HEPOXILIN B3	0.01	-0.03	-0.01	-0.05	0.00	0.00	0.02	0.63
8	Lip1	D04	5(S),12(R)-DIHETE all trans	0.02	-0.02	-0.04	-0.04	0.01	0.00	0.01	0.75
9	Lip1	D10	11,12-EPOXYEICOSATRIENOIC ACID	0.00	-0.04	-0.02	-0.04	0.02	0.00	0.02	0.52
10	Lip1	H03	5-KETOEICOSATETRAENOIC ACID	0.04	-0.02	-0.05	-0.04	0.00	0.01	0.01	1.41
11	Lip2	F09	1-STEAROYL-2-ARACHIDONOYL-GLYCEROL	0.01	-0.03	-0.05	-0.05	-0.01	0.01	-0.01	0.45
12	Lip3	B01	EICOSAPENTAENOIC ACID (20:5 n-3)	0.07	0.01	-0.04	-0.07	0.03	0.03	0.05	1.43
13	Lip3	B03	ARACHIDONIC ACID (20:4 n-6)	0.01	-0.05	-0.05	-0.03	0.02	0.01	0.03	1.39
14	Lip3	F09	2-ARACHIDONOYLGLYCEROL	0.00	-0.03	-0.05	-0.05	0.00	0.01	0.02	0.53
*15	5039	C05	Dipyridamole	0.38	0.53	-0.04	0.09	0.39	0.25	0.11	0.53

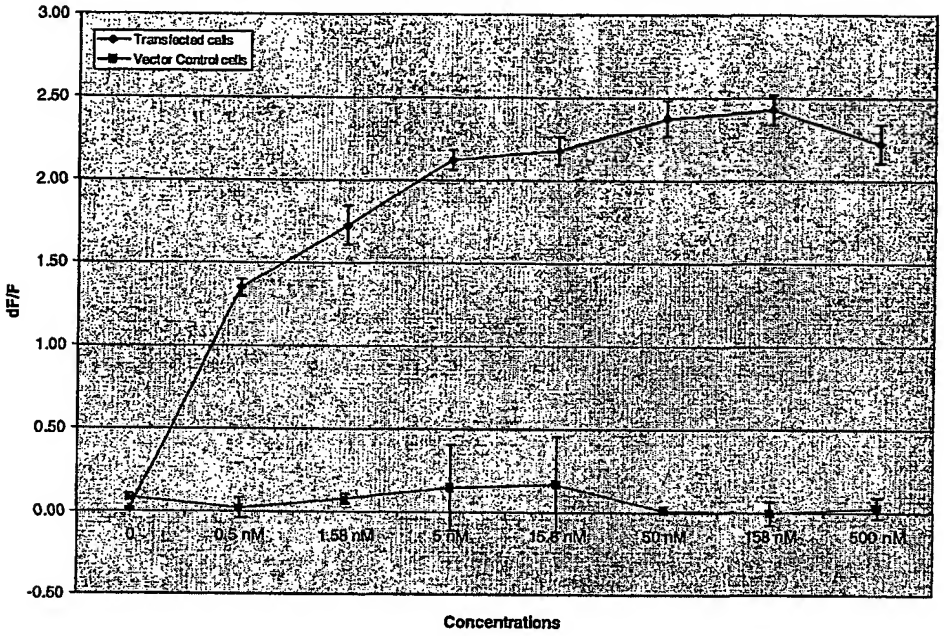
After priming, 15 compounds show a Ca^{2+} transient signal. 14 of them come from the lipids collection. These compounds are also tested against wild type cells and other transfected cells in order to check the selectivity. The 5-KETOEICOSATETRAENOIC ACID shows a dF/F = 1.41 and it is not active on the non-transfected cells.

Validation

Validation consists in testing the compounds in a dose response manner in order to evaluate their EC50. Each compound is tested at 8 concentrations in quadruplicates. The stock concentrations of 5-KETOEICOSATETRAENOIC ACID is 0.1 mM. The dose response range tested is the following : 500 nM, 158 nM, 50 nM, 15.8 nM, 5 nM, 1.58 nM, 0.5 nM and 0.

Cell plates preparation is the same as in ligand fishing (see above). Cells are primed with ATP 10 µM and incubated for 30 minutes before compound injection.

Fig. 1 : 5-KETOEICOSATETRAENOIC DOSE RESPONSE CURVE



The 5-KETOEICOSATETRAENOIC ACID EC50 is 0.5 nM.

cAMP assay

For measurements of adenylyl cyclase activity, cells are seeded in 24 well plates at a density of 100000 cells per well and maintained in culture for 2-3 days. On the day of the assay, cells are washed and incubated for 4h in serum-free medium containing [3H]adenine (Amersham; 2uCi/ml), to label the intracellular ATP pool. Where required, pertussis toxin (PTX; Sigma, 100 ng/ml) is included during this labelling period.

Thereafter, plates are washed and cells are incubated in 500 ul of HBS (130 mM NaCl, 0.8 mM MgSO4, 5.4 mM KCL, 0.9 mM NaH2PO4, 1.8 mM CaCl2, 25 mM glucose, 20 mM HEPES, pH 7.4) supplemented with isobutylmethylxanthine (IBMX; Sigma, 1mM) at 37ºC. Cells are then stimulated with the diterpene forskolin (FSK; Sigma, 3 uM), and receptor agonists in appropriate concentrations. Under these assay conditions receptors coupled either positively or negatively to AC can be detected as they increase or decrease the FSK-stimulated signal, respectively.

Following 15 min of incubation, the reactions are stopped by aspiration of the buffer and cell extraction with 1 ml of ice-cold trichloroacetic acid (5%). Pools of [3H]ATP and [3H]cAMP are separated by sequential chromatography on Dowex and alumina columns.

Claims

1. An isolated polypeptide selected from one of the groups consisting of:
(a) an isolated polypeptide encoded by a polynucleotide comprising the sequence of SEQ ID NO 1 or SEQ ID No.5;
(b) an isolated polypeptide comprising a polypeptide sequence having at least 95% identity to the polypeptide sequence of SEQ ID NO 2 or SEQ ID NO.6;
(c) an isolated polypeptide having at least 95% identity to the plypeptide sequence of SEQ ID NO 2 or SEQ ID NO.6; and
(d) the polypeptide sequence of SEQ ID NO 2 or SEQ ID NO.6 and
(e)fragments and variants of such polypeptides in (a) to (d).

2. The isolated polypeptide as claimed in claim 1 comprising the polypeptide sequence of SEQ ID NO 2 or SEQ ID NO.6.

3. An isolated polynucleotide selected from one of the groups consisting of:
(a) an isolated polynucleotide comprising a polynucleotide sequence having at least 95% identity to the polynucleotide sequence of SEQ ID NO 1 or SEQ ID NO.5;
(b) an isolated polynucleatide having at least 95% identity to the polynucleotide of SEQ ID NO 1 or SEQ ID NO 5;
(c) an isolated polynucleotide comprising a polynucleotide sequence encoding a polypeptide sequence having at least 95% identity to the polypeptide sequence of SEQ ID NO 2 or SEQ ID NO.6;
(d) an isolated polynucleotide having a polynucleotide sequence encoding a polypeptide sequence having at least 95% identity to the polypeptide sequence of SEQ ID NO 2 or SEQ ID NO.6;
(e) an isolated polynucleotide with a nucleotide sequence of at least 100 nucleotides obtained by screening a library under stringent hybridization conditions with a labeled probe having the sequence of SEQ ID NO 1 or SEQ ID NO 5 or a fragment thereof having at least 15 nucleotides;
(f) a polynucleotide which is the RNA equivalent of a polynucleotide of (a) to (e); or a polynucleotide sequence complementary to said isolated polynucleotide and polynucleotides that are variants and fragments of the above mentioned polynucleotides or that are complementary to above mentioned polynucleotides, over the entire length thereof.

4. An isolated polynucleotide as claimed in claim 3 selected from the group consisting of:
(a) an isolated polynucleotide comprising the polynucleotide of SEQ ID NO 1 or SEQ ID NO 5 ;
(b) the isolated polynucleotide of SEQ ID NO 1 or SEQ ID NO 5;
(c) an isolated polynucleotide comprising a polynucleotide sequence encoding the polypeptide of SEQ ID NO 2 or SEQ ID NO 6; and
(d) an isolated polynucleotide encoding the polypeptide of SEQ ID NO 2 or SEQ ID NO 6.

5. An expression system comprising a polynucleotide capable of producing a polypeptide of claim 1 when said expression vector is present in a compatible host cell.

6. A recombinant host cell comprising the expression vector of claim 5 or a membrane thereof expressing the polypeptide of claim 1.

7. A process for producing a polypeptide of claim 1 comprising the step of culturing a host cell as defined in claim 6 under conditions sufficient for the production of said polypeptide and recovering the polypeptide from the culture medium.

8. A fusion protein consisting of the Immunoglobulin Fc-region and any one polypeptide of claim 1.

9. An antibody immunospecific for the polypeptide of any one of claims 1 to 3.

10. A method for screening to identify compounds that stimulate or inhibit the function or level of the polypeptide of claim 1 comprising a method selected from the group consisting of:
(a) measuring or, detecting, quantitatively or qualitatively, the binding of a candidate compound to the polypeptide (or to the cells or membranes expressing the polypeptide) or a fusion protein thereof by means of a label directly or indirectly associated with the candidate compound;
(b) measuring the competition of binding of a candidate compound to the polypeptide (or to the cells or membranes expressing the polypeptide) or a fusion protein thereof in the presence of a labeled competitor;
(c) testing whether the candidate compound results in a signal generated by activation or inhibition of the polypeptide, using detection systems appropriate to the cells or cell membranes expressing the polypeptide;

(d) mixing a candidate compound with a solution containing a polypeptide of claim 1, to form a mixture, measuring activity of the polypeptide in the mixture, and comparing the activity of the mixture to a control mixture which contains no candidate compound; or
(e) detecting the effect of a candidate compound on the production of mRNA encoding said polypeptide or said polypeptide in cells, using for instance, an ELISA assay, and
(f) producing said compound according to biotechnological or chemical standard techniques.

SEQUENCE LISTING

<110> Novartis AG

<120> Novel G Protein-Coupled Receptors and DNA Sequences Thereof

<130> 4-32004

<160> 8

<170> PatentIn version 3.0

<210> 1

<211> 1155

<212> DNA

<213> Homo sapiens

<400> 1

atggaacttc ataacctgag ctctccatct ccctctctct cctcctctgt tctccctccc 60

tccttctctc cctcaccctc ctctgctccc tctgccttta ccactgtggg ggggtcctct 120

ggagggccct gccaccccac ctcttcctcg ctggtgtctg ccttcctggc accaatcctg 180

gccctggagt ttgtcctggg cctggtgggg aacagtttgg ccctcttcat cttctgcatc 240

cacacgcggc cctggacctc caacacggtg ttcctggtca gcctggtggc cgctgacttc 300

ctcctgatca gcaacctgcc cctccgcgtg gactactacc tcctccatga gacctggcgc 360

tttggggctg ctgcctgcaa agtcaacctc ttcatgctgt ccaccaaccg cacggccagc 420

gttgtcttcc tcacagccat cgcactcaac cgctacctga aggtggtgca gccccaccac 480

gtgctgagcc gtgcttccgt gggggcagct gcccgggtgg ccgggggact ctgggtgggc 540

atcctgctcc tcaacgggca cctgctcctg agcaccttct ccggcccctc ctgcctcagc 600

tacagggtgg gcacgaagcc ctcggcctcg ctccgctggc accaggcact gtacctgctg 660

gagttcttcc tgccactggc gctcatcctc tttgctattg tgagcattgg gctcaccatc 720

cggaaccgtg gtctgggcgg gcaggcaggc ccgcagaggg ccatgcgtgt gctggccatg 780

gtggtggccg tctacaccat ctgcttcttg cccagcatca tctttggcat ggcttccatg 840

gtggctttct ggctgtccgc ctgccgatcc ctggacctct gcacacagct cttccatggc 900

tccctggcct tcacctacct caacagtgtc ctggaccccg tgctctactg cttctctagc 960

cccaacttcc tccaccagag ccgggccttg ctgggcctca cgcggggccg gcagggccca 1020

gtgagcgacg agagctccta ccaaccctcc aggcagtggc gctaccggga ggcctctagg 1080

aaggcggagg ccatagggaa gctgaaagtg cagggcgagg tctctctgga aaaggaaggc 1140

tcctcccagg gctga 1155

<210> 2

<211> 384

<212> PRT

<213> Homo sapies

<400> 2

```
Met Glu Leu His Asn Leu Ser Ser Pro Ser Pro Ser Leu Ser Ser Ser
1               5                  10                  15

Val Leu Pro Pro Ser Phe Ser Pro Ser Pro Ser Ser Ala Pro Ser Ala
            20                  25                  30

Phe Thr Thr Val Gly Gly Ser Ser Gly Gly Pro Cys His Pro Thr Ser
        35                  40                  45

Ser Ser Leu Val Ser Ala Phe Leu Ala Pro Ile Leu Ala Leu Glu Phe
    50                  55                  60

Val Leu Gly Leu Val Gly Asn Ser Leu Ala Leu Phe Ile Phe Cys Ile
65                  70                  75                  80

His Thr Arg Pro Trp Thr Ser Asn Thr Val Phe Leu Val Ser Leu Val
                85                  90                  95

Ala Ala Asp Phe Leu Leu Ile Ser Asn Leu Pro Leu Arg Val Asp Tyr
            100                 105                 110

Tyr Leu Leu His Glu Thr Trp Arg Phe Gly Ala Ala Ala Cys Lys Val
        115                 120                 125

Asn Leu Phe Met Leu Ser Thr Asn Arg Thr Ala Ser Val Val Phe Leu
    130                 135                 140

Thr Ala Ile Ala Leu Asn Arg Tyr Leu Lys Val Val Gln Pro His His
145                 150                 155                 160

Val Leu Ser Arg Ala Ser Val Gly Ala Ala Ala Arg Val Ala Gly Gly
                165                 170                 175

Leu Trp Val Gly Ile Leu Leu Leu Asn Gly His Leu Leu Leu Ser Thr
            180                 185                 190

Phe Ser Gly Pro Ser Cys Leu Ser Tyr Arg Val Gly Thr Lys Pro Ser
        195                 200                 205

Ala Ser Leu Arg Trp His Gln Ala Leu Tyr Leu Leu Glu Phe Phe Leu
    210                 215                 220

Pro Leu Ala Leu Ile Leu Phe Ala Ile Val Ser Ile Gly Leu Thr Ile
225                 230                 235                 240

Arg Asn Arg Gly Leu Gly Gly Gln Ala Gly Pro Gln Arg Ala Met Arg
                245                 250                 255

Val Leu Ala Met Val Val Ala Val Tyr Thr Ile Cys Phe Leu Pro Ser
            260                 265                 270

Ile Ile Phe Gly Met Ala Ser Met Val Ala Phe Trp Leu Ser Ala Cys
        275                 280                 285

Arg Ser Leu Asp Leu Cys Thr Gln Leu Phe His Gly Ser Leu Ala Phe
```

290 295 300

Thr Tyr Leu Asn Ser Val Leu Asp Pro Val Leu Tyr Cys Phe Ser Ser
305 310 315 320

Pro Asn Phe Leu His Gln Ser Arg Ala Leu Leu Gly Leu Thr Arg Gly
325 330 335

Arg Gln Gly Pro Val Ser Asp Glu Ser Ser Tyr Gln Pro Ser Arg Gln
340 345 350

Trp Arg Tyr Arg Glu Ala Ser Arg Lys Ala Glu Ala Ile Gly Lys Leu
355 360 365

Lys Val Gln Gly Glu Val Ser Leu Glu Lys Glu Gly Ser Ser Gln Gly
370 375 380

<210> 3

<211> 23

<212> DNA

<213> Homo sapiens

<400> 3
tcttcatctt ctgcatccac acg 23

<210> 4

<211> 23

<212> DNA

<213> Homo sapiens

<400> 4
agtggcagga agaactccag cag 23

<210> 5

<211> 1041

<212> DNA

<213> Homo sapiens

<400> 5
atgtacaacg ggtcgtgctg ccgcatcgag ggggacacca tctcccaggt gatgccgccg 60

ctgctcattg tggcctttgt gctgggcgca ctaggcaatg gggtcgccct gtgtggtttc 120

tgcttccaca tgaagacctg gaagcccagc actgtttacc ttttcaattt ggccgtggct 180

gatttcctcc ttatgatctg cctgcctttt cggacagact attacctcag acgtagacac 240

tgggcttttg gggacattcc ctgccgagtg gggctcttca cgttggccat gaacagggcc 300

gggagcatcg tgttccttac ggtggtggct gcggacaggt atttcaaagt ggtccacccc 360

caccacgcgg tgaacactat ctccacccgg gtggcggctg gcatcgtctg caccctgtgg 420

gccctggtca tcctgggaac agtgtatctt ttgctggaga accatctctg cgtgcaagag 480

acggccgtct cctgtgagag cttcatcatg gagtcggcca atggctggca tgacatcatg 540

ttccagctgg agttctttat gcccctcggc atcatcttat tttgctcctt caagattgtt 600

tggagcctga ggcggaggca gcagctggcc agacaggctc ggatgaagaa ggcgacccgg 660

ttcatcatgg tggtggcaat tgtgttcatc acatgctacc tgcccagcgt gtctgctaga 720

ctctatttcc tctggacggt gccctcgagt gcctgcgatc cctctgtcca tggggccctg 780

cacataaccc tcagcttcac ctacatgaac agcatgctgg atcccctggt gtattatttt 840

tcaagccccc cctttcccaa attctacaac aagctcaaaa tctgcagtct gaaacccaag 900

cagccaggac actcaaaaac acaaaggccg gaagagatgc caatttcgaa cctcggtcgc 960

aggagttgca tcagtgtggc aaatagtttc caaagccagt ctgatgggca atgggatccc 1020

cacattgttg agtggcactg a 1041

<210> 6

<211> 346

<212> PRT

<213> Homo sapiens

<400> 6

Met Tyr Asn Gly Ser Cys Cys Arg Ile Glu Gly Asp Thr Ile Ser Gln
1 5 10 15

Val Met Pro Pro Leu Leu Ile Val Ala Phe Val Leu Gly Ala Leu Gly
20 25 30

Asn Gly Val Ala Leu Cys Gly Phe Cys Phe His Met Lys Thr Trp Lys
35 40 45

Pro Ser Thr Val Tyr Leu Phe Asn Leu Ala Val Ala Asp Phe Leu Leu
50 55 60

Met Ile Cys Leu Pro Phe Arg Thr Asp Tyr Tyr Leu Arg Arg Arg His
65 70 75 80

Trp Ala Phe Gly Asp Ile Pro Cys Arg Val Gly Leu Phe Thr Leu Ala
85 90 95

Met Asn Arg Ala Gly Ser Ile Val Phe Leu Thr Val Val Ala Ala Asp
100 105 110

Arg Tyr Phe Lys Val Val His Pro His His Ala Val Asn Thr Ile Ser
115 120 125

Thr Arg Val Ala Ala Gly Ile Val Cys Thr Leu Trp Ala Leu Val Ile

130 135 140

Leu Gly Thr Val Tyr Leu Leu Leu Glu Asn His Leu Cys Val Gln Glu
145 150 155 160

Thr Ala Val Ser Cys Glu Ser Phe Ile Met Glu Ser Ala Asn Gly Trp
165 170 175

His Asp Ile Met Phe Gln Leu Glu Phe Phe Met Pro Leu Gly Ile Ile
180 185 190

Leu Phe Cys Ser Phe Lys Ile Val Trp Ser Leu Arg Arg Arg Gln Gln
195 200 205

Leu Ala Arg Gln Ala Arg Met Lys Lys Ala Thr Arg Phe Ile Met Val
210 215 220

Val Ala Ile Val Phe Ile Thr Cys Tyr Leu Pro Ser Val Ser Ala Arg
225 230 235 240

Leu Tyr Phe Leu Trp Thr Val Pro Ser Ser Ala Cys Asp Pro Ser Val
245 250 255

His Gly Ala Leu His Ile Thr Leu Ser Phe Thr Tyr Met Asn Ser Met
260 265 270

Leu Asp Pro Leu Val Tyr Tyr Phe Ser Ser Pro Ser Phe Pro Lys Phe
275 280 285

Tyr Asn Lys Leu Lys Ile Cys Ser Leu Lys Pro Lys Gln Pro Gly His
290 295 300

Ser Lys Thr Gln Arg Pro Glu Glu Met Pro Ile Ser Asn Leu Gly Arg
305 310 315 320

Arg Ser Cys Ile Ser Val Ala Asn Ser Phe Gln Ser Gln Ser Asp Gly
325 330 335

Gln Trp Asp Pro His Ile Val Glu Trp His
340 345

<210> 7

<211> 23

<212> DNA

<213> Homo sapiens

<400> 7
ctcttcacgt tggccatgaa cag 23

<210> 8

<211> 23

<212> DNA

<213> Homo sapiens

<400> 8

ggcaccgtcc agaggaaata gag 23